Prompting strategy

LLM-based extraction of triage-note variables was performed using a zero-shot prompting approach. The system prompt instructed the model to act as an expert in medical documentation analysis and to classify each predefined variable as either True or False. If a variable was not explicitly mentioned in the triage note, it was classified as False.

System prompt:
You are an expert in analyzing medical documentation. You will be provided an Emergency Department triage note delineated by <triage-note>, and your task is to extract structured data by classifying each parameter as either True or False. If there is insufficient information, classify the variable as False.

Extraction targets:

- Was the patient referred to the ED by another provider?
- Were any interventions performed prior to arrival?
- Was the blood sugar reported to be under 60 mg/dL?
- Was the blood sugar reported to be over 300 mg/dL?
- Were antibiotics given prior to arrival?
- Was intoxication mentioned?
- Was cognitive impairment mentioned?
- Was recent surgery mentioned?
- Was immunocompromised status mentioned?
- Was dialysis status mentioned?

Structured output and schema constraints

To ensure consistent and reproducible outputs, LLM responses were constrained to a predefined JSON schema using the Outlines library. Each variable was returned as a binary True/False value, and free-text responses were disallowed. This approach minimized output variability and ensured deterministic extraction behavior.

Reproducibility and code availability

The full prompting logic, JSON schema, and inference pipeline used for LLM-based feature extraction are publicly available at:
https://github.com/fzeinali/ED-disposition-LLM-triage-notes